# Virtual Machine Resource Monitoring and Networking of Virtual Machines

Ananth I. Sundararaj

Department of Computer Science
Northwestern University
July 07, 2003

# Outline

- Efficient Monitoring of Virtual Machine Resources
  - Objective
  - Motivation
  - Basic Approach
  - Experimental Setup
  - Research Issues
  - Results and Discussion
  - Conclusions
  - Future Work

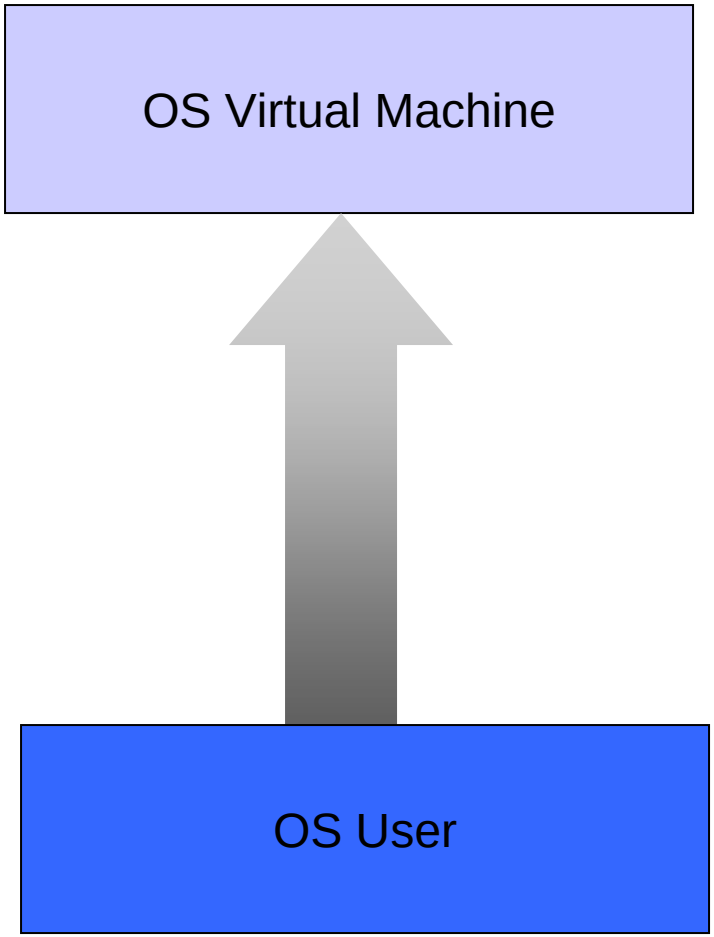
# Objective

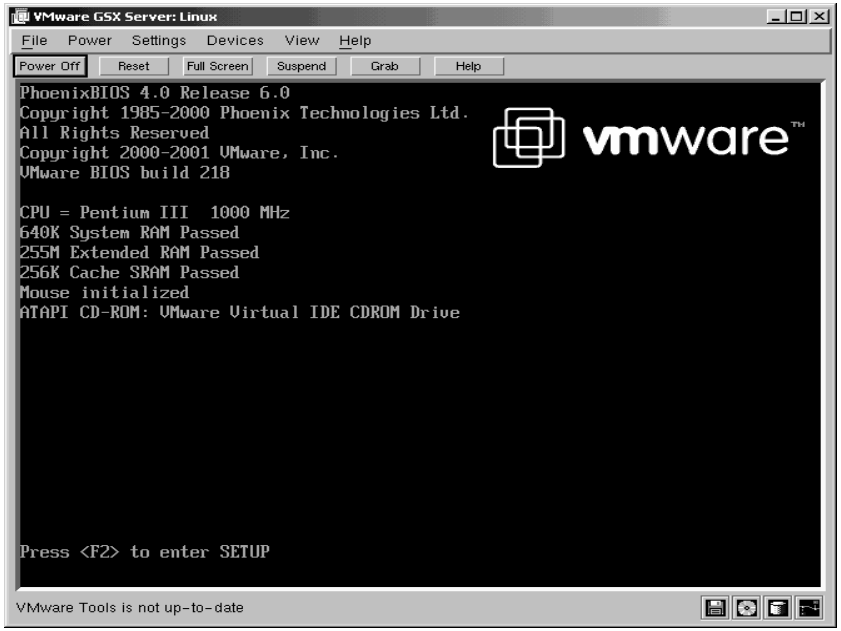
- Problem Statement
  - To address the problem of efficient monitoring of virtual machine resources hosted on a physical host machine
  - Given the monitoring information in the host operating system, attempt to reconstruct the monitoring information in the guest operating system residing on the virtual machines
  - To characterize the aggregate system performance using time series analysis
  - To develop a mapping from aggregate system resources to individual virtual machine system resources

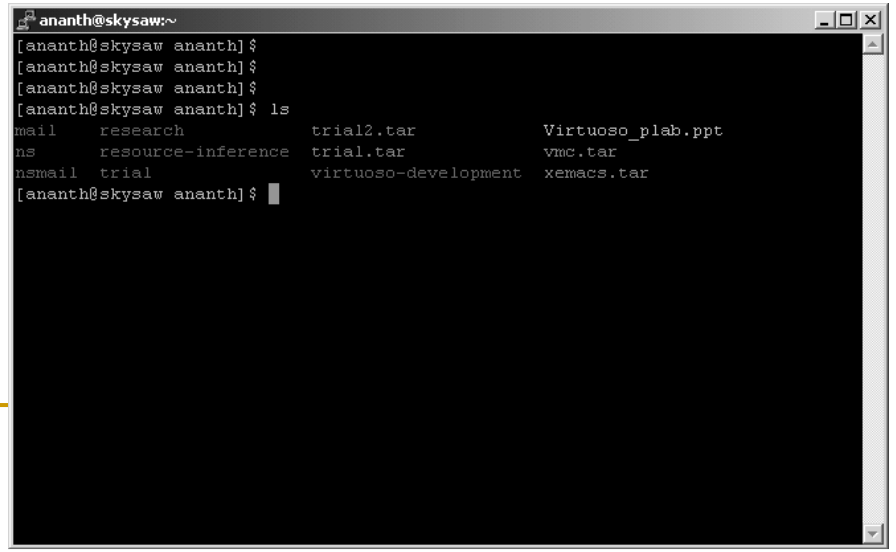
# Motivation

- Abstraction of a Virtual Machine
- Research areas and projects where this abstraction is being leveraged
- Why is the problem important
- Need for efficient monitoring

# Abstraction of a Virtual Machine



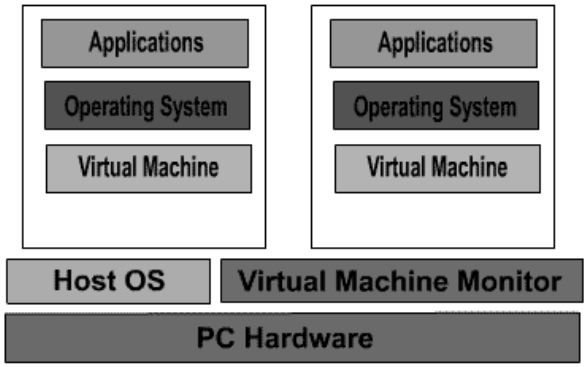



# Virtual Machine

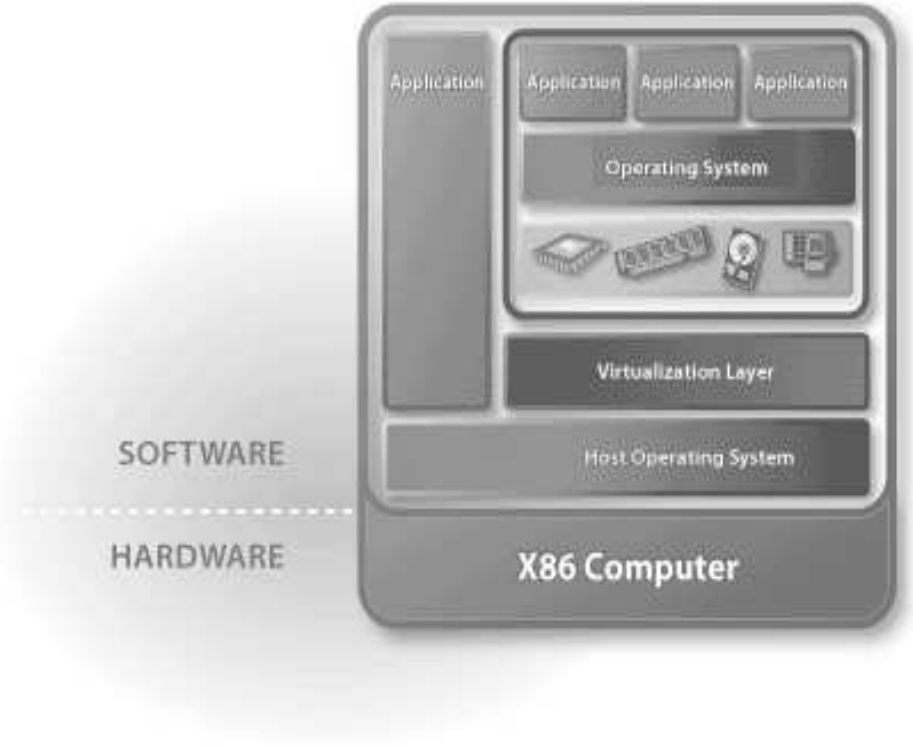
- History
  - First came about in the 1960's on mainframes as a way to create less complex multi user time share environments
- What is it?
  - A virtual machine is an abstraction of a physical machine
    - Created using a Virtual Machine Monitor (VMM) running on a physical machine
  - Gives the illusion of working on a separate machine

# Virtual machines contd..

- Architecture
  - The abstraction of a virtual machine is that each user appears to have a dedicated machine at their disposal, the hardware of which they can access directly




Hosted Virtual Machine

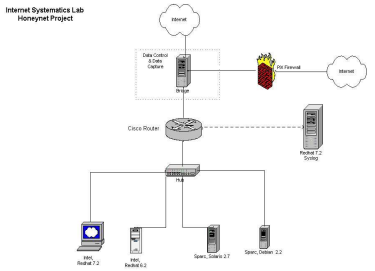


# Areas where this abstraction is being leveraged

- Grid Computing on Virtual Machines

- Prototyping

- Virtual Honeynets used as a counter intrusion strategy



# Outline

- Objective
- Motivation
- Basic Approach
- Experimental Setup
- Research Issues
- Results and Discussion
- Conclusions
- Future Work

# Basic Approach

- Typical monitoring system on a physical machine
- Aggregate system performance is characterized using time series analysis
- A mapping from aggregate system resources to individual virtual machine resources is developed
- Model developed could then be used to build monitoring tools for such systems

# Experimental Setup

- Physical machine is a dual Pentium III/800 MHz with 1 GB memory running RedHat 7.1
- Virtual machine uses VMware GSX server with 128 MB memory and RedHat 7.3
- Case I
  - A physical machine hosts a single virtual machine
- Case II
  - A physical machine hosts two virtual machines

# Data Collection

- Time synchronization
- Reading data from /proc of physical and virtual machine
  - Tool written by Luka Spoljaric
  - Typical usage: bash$ ns [-max i] [-rate f] [-period f] [-name s] [-timestamp]
- Counters read
  - CPU
    - Load
    - Number of processes
    - Usage
    - Context Switches
  - Memory
    - Page faults
    - % usage of buffer
  - Disk
    - Bytes transferred (read and write operations)
  - Network
    - Bytes transferred (transmitted and received)

# Possible Scenarios

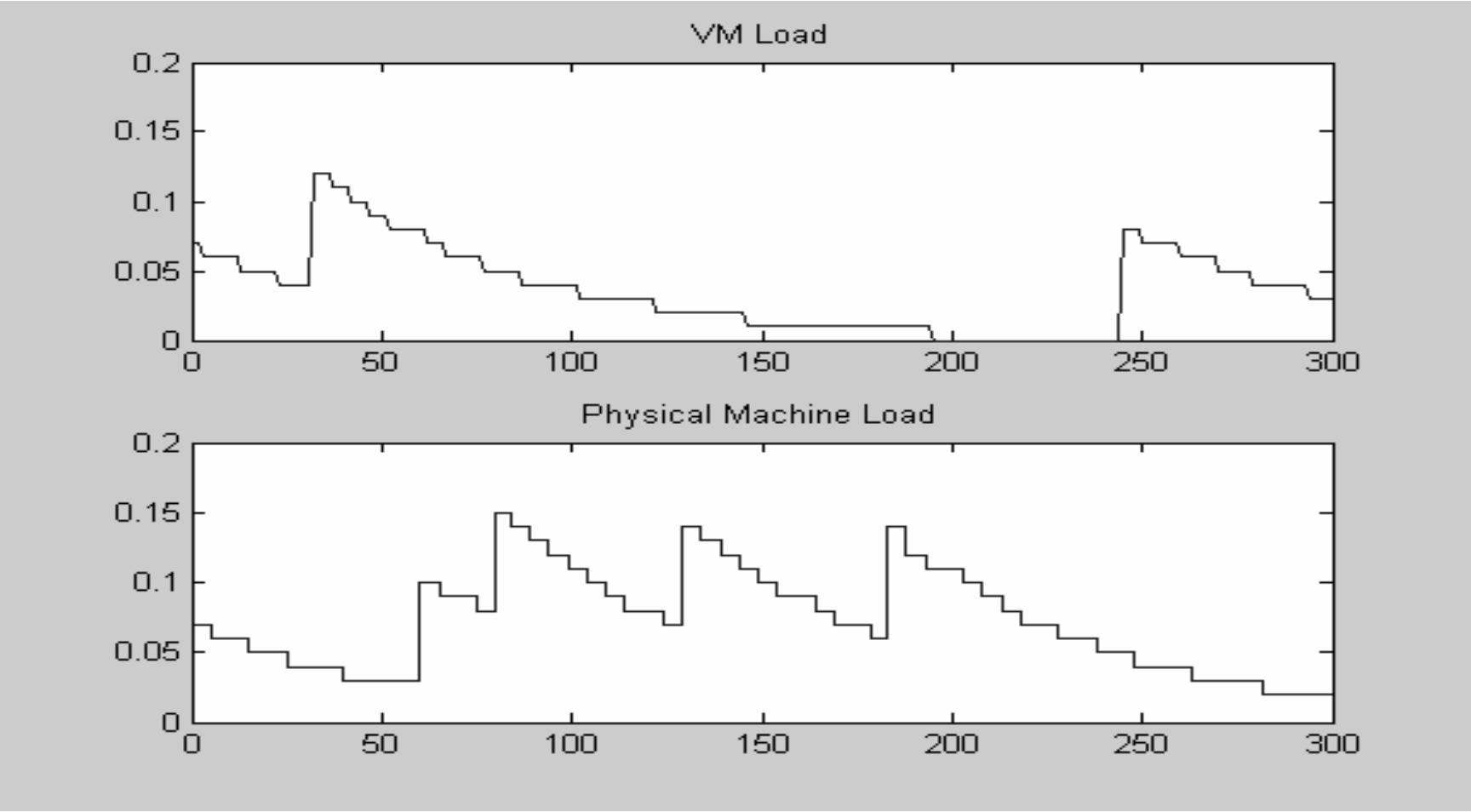
| Physical Machine | Virtual Machine |
|---|---|
| Completely Unloaded | - |
| Only load process | - |
| Only virtual machine | Completely unloaded |
| Virtual machine + load process | Completely unloaded |
| Virtual machine + VM load process | VM load process |
| Virtual machine + load process + VM load | VM load process |

Load Processes: The background load was produced by host load trace playback

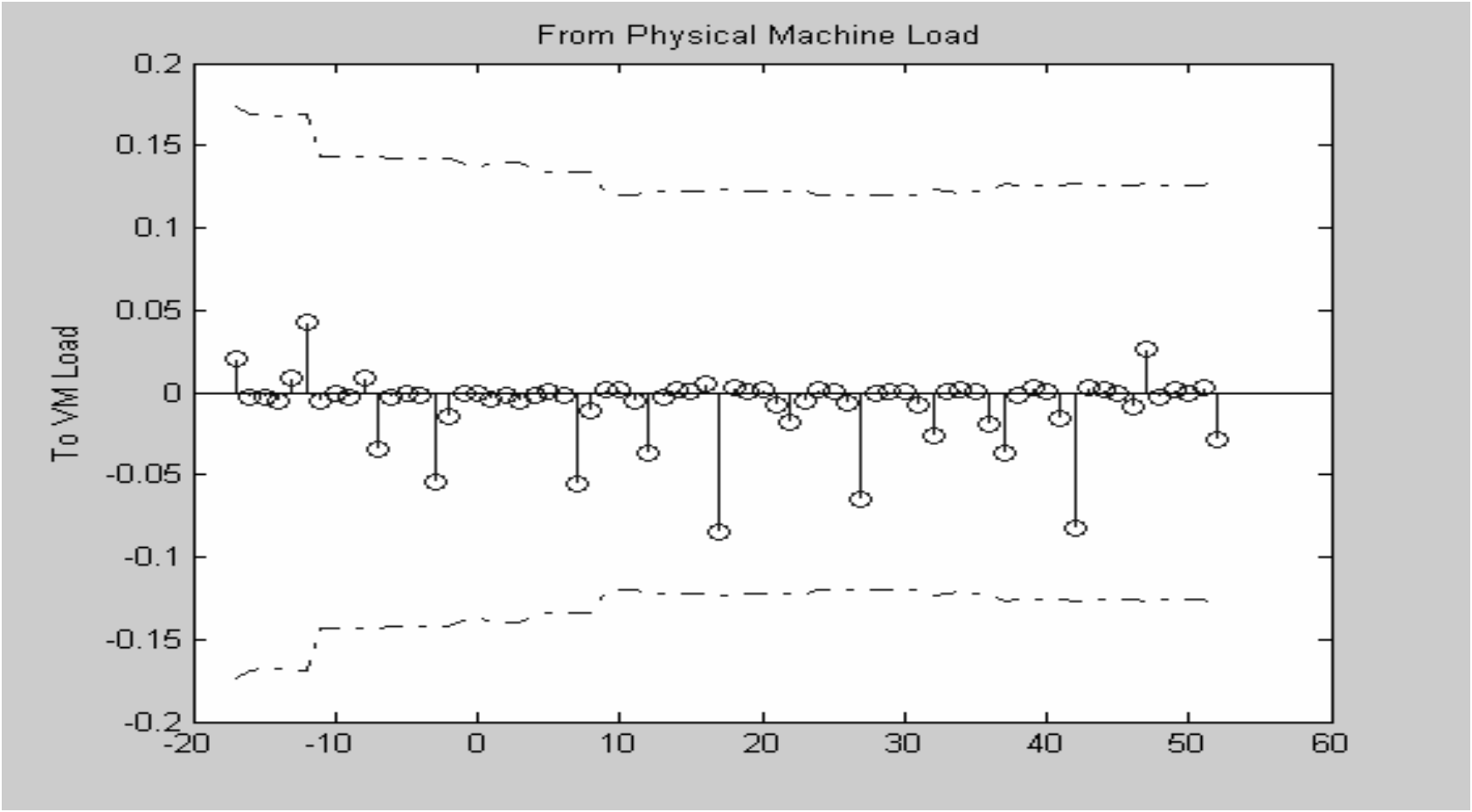
# Research Issues

- Effect of load process in physical machine on load in virtual machine
- Rate of execution in the Virtual Machine
- Multiple input single output analysis
- Other benchmarks
- Alternatives to reading /proc
- Analysis from the view of virtual machine as a process

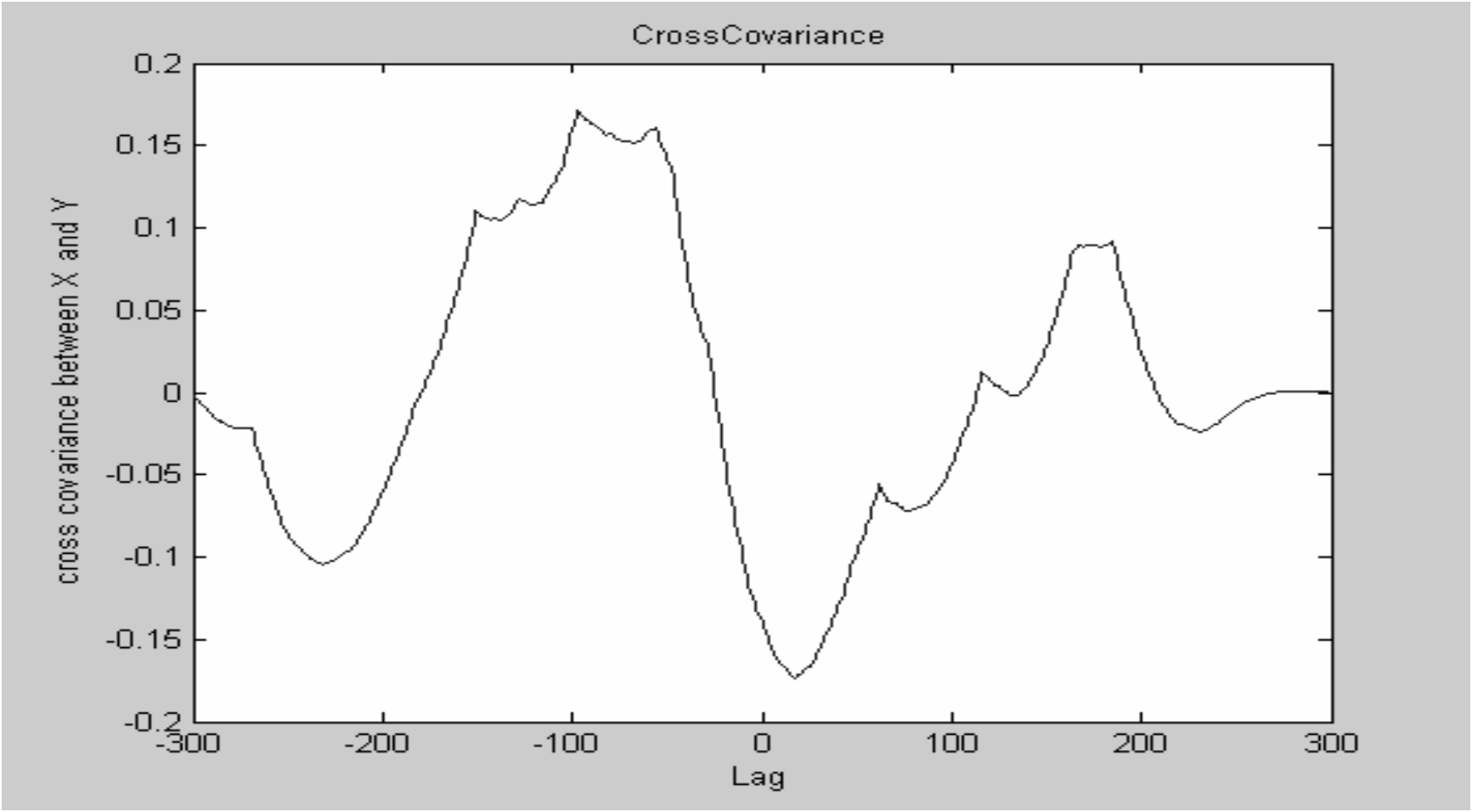
# Results and Discussion



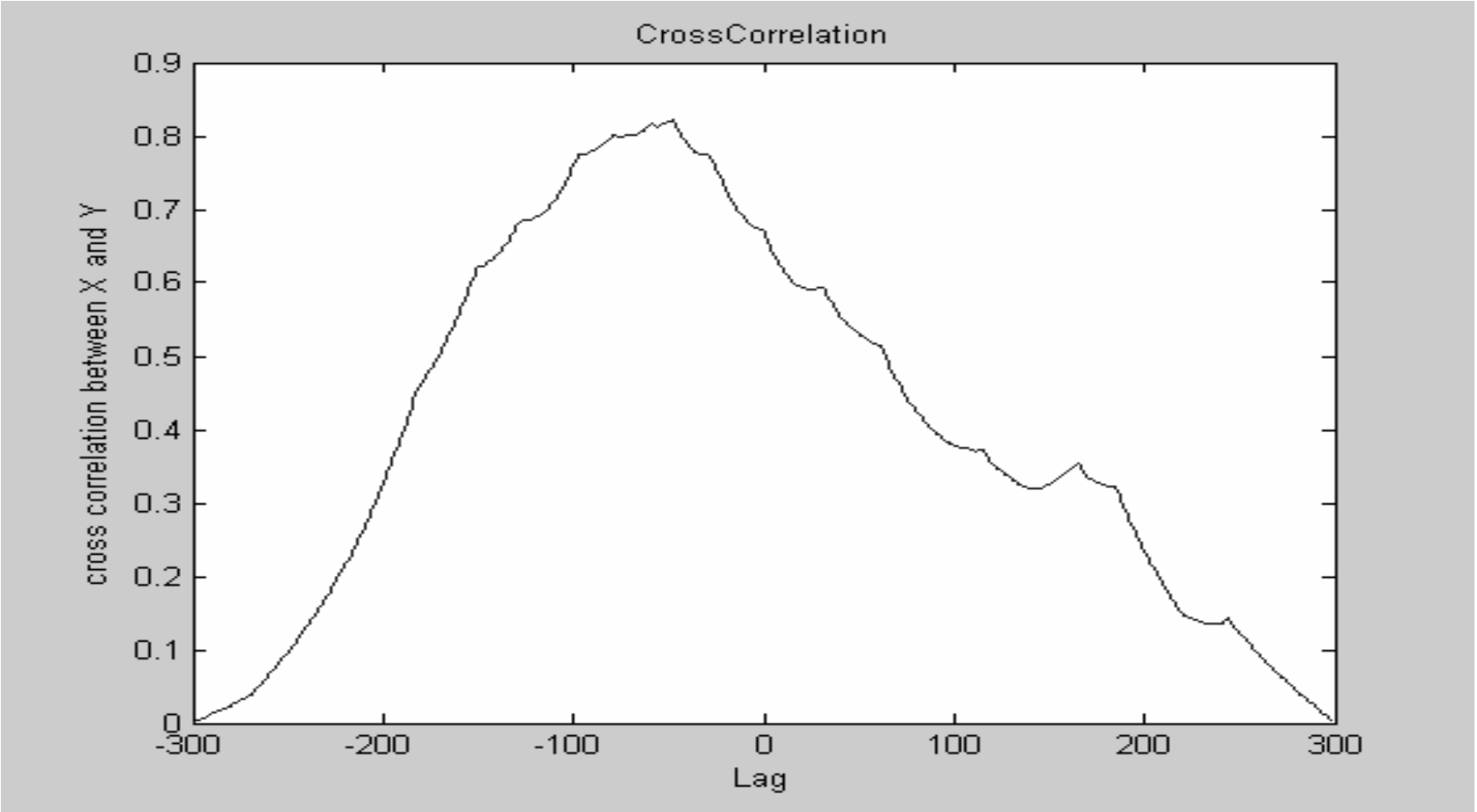
# Impulse Response Function



# Cross Covariance



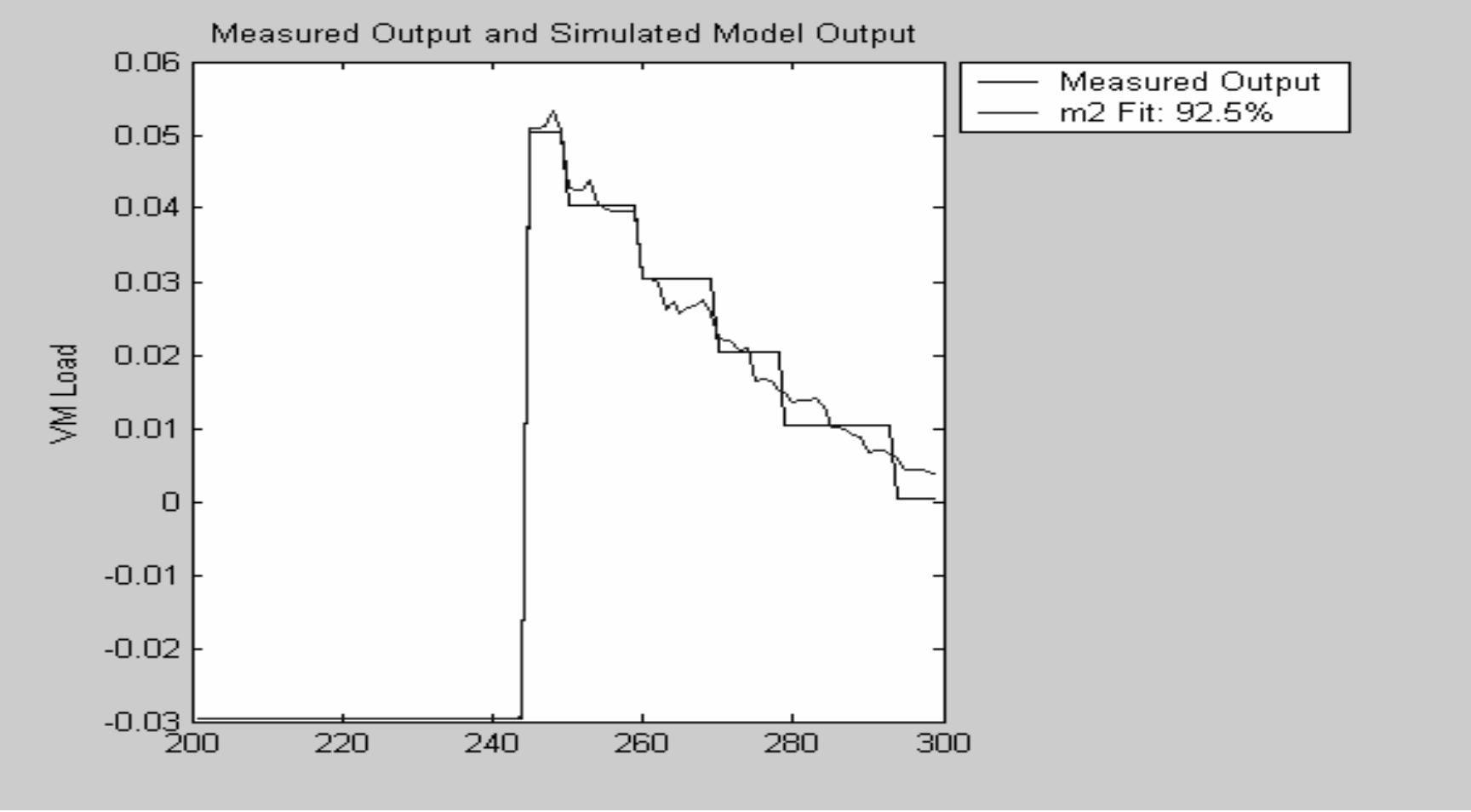
# Cross Correlation



# Fitted Model

- Basic Dynamic Model
  - The basic relationship is the linear difference equation
  - ARX Model
    - General form is
      - $y(t) + a_1 y(t-T) + a_2 t(t-2T) = b_1 u(t-2T) + b_2 u(t-3T) + e(t)$
    - Parameters (20, 17, 50) (poles, zeros, delay)

# Model Validation



# Outline

- Objective
- Motivation
- Basic Approach
- Experimental Setup
- Research Issues
- Results and Discussion
- Conclusions
- Future Work

# Conclusions

- Provided motivation for efficient monitoring of virtual machines hosted on physical machines
- Detailed the approach adopted
- Described the experimental setup
- Discussed the preliminary results

# Future Work

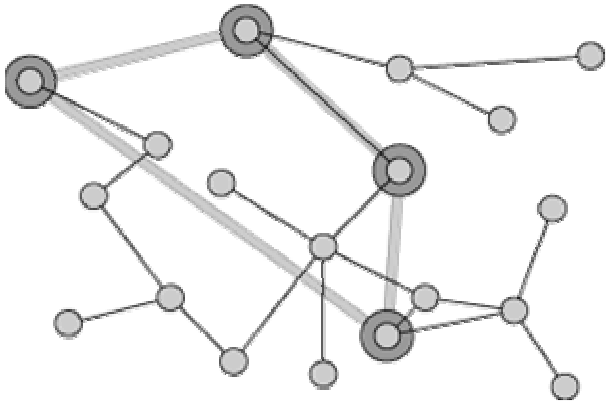
- To come up with a more generic model considering all the cases and scenarios listed
- To collect data differently and perhaps apply different analysis techniques
- Based on the models developed to build monitoring tools for systems hosting many virtual machines on a single physical host

# Outline – Current Work

- Network of Virtual Machines
  - Scenario
  - Objectives
  - Problem Formulation
  - Issues

# Scenario

- Virtual Machine Networking
  - Scenario

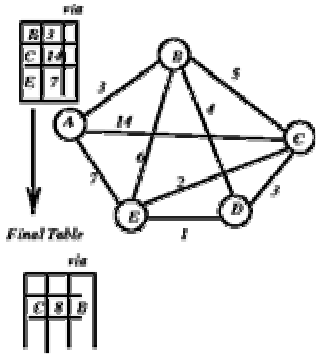


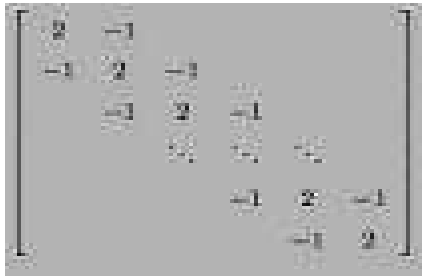
# Objectives

- An overlay network could be formed among the remote virtual machines giving rise to a virtual LAN
- The overlay network could optimize itself with respect to the communication between the virtual machines
- To maintain network connectivity during and after migration of virtual machines

# Abstract Problem Formulation

- Network organization and management as a state machine
- Concept of a state for a network
  - Topology
  - Routing information
- The inputs to the state machine
  - Bandwidth matrix
  - Latency matrix



# Issues Involved

- Collecting network and topology information
- Inferring current state
- Generating inputs
- Dynamically changing state